PROBA-V PERFORMANCE ASSESSMENT FOR FOREST COVER MAPPING OVER THE ATLANTIC BIOGEOGRAPHICAL REGION OF EUROPE

Assia Azzi[1], Ioannis Manakos[1], Chariton Kalaitzidis[1], Mathias Schardt[2], and Heinz Gallaun[2]

1. Mediterranean Agronomic Institute of Chania, Department of Geoinformation in Environmental Management, Chania, Greece; azziassia(at)yahoo.fr, ioannis.manakos(at)earsel.org, chariton(at)maich.gr
2. JOANNEUM REASERCH Forschungsgesellschaft mbH, Remote Sensing and Geoinformation, Graz, Austria; mathias.schardt(at)joanneum.at

ABSTRACT

The Belgian Federal Science Policy Office (BELSPO) has initiated a Preparatory Evaluation/Validation Programme for the products of the new PROBA-V satellite to be launched in 2012. The satellite will allow daily monitoring of terrestrial vegetation cover through remote sensing, and will cover the data provision gap between the closure of the SPOT/VEGETATION Programme and the launch of the SENTINEL-3 mission. The aim of this study is to evaluate the improvements that PROBA-V will bring along for forest monitoring in the Atlantic Biogeographical Region of Europe, and lies within the objectives of the FM@PROBA-V project. A representative site in Northern Portugal is selected for this reason. VEGETATION, LANDSAT-TM5, and MODIS data along with the JRC Forest Cover Map are used to train the classifiers, simulate PROBA-V data, apply the classifiers at 250 m, 1/3 of a km, and 1 km pixels, and validate the results, while quantifying the accuracies. Maximum Likelihood (ML), Artificial Neural Networks (ANN), and Support Vector Machine (SVM) methods were tested. From the confusion matrices the best result is obtained by MODIS 2 bands with ANN classifier. Further analysis on the base of those confusion matrices will be applied to define the best classifier taking into account all the parameters of the matrices. The best performing classifier will then be recommended to examine its robustness against sudden disastrous events, like fire, in the same area, performing change detection between sequential dates (before and after the event). The performance of the data and classifiers are demonstrated, and the preliminary results are discussed.

INTRODUCTION

World forests cover roughly 31% of land area and in the last decade it has been reduced at a rate of 13 million hectares per year (1). According to the global forest resources assessment 2010 of the Food and Agriculture Organization, significant progress has been made towards reversing the overall trend of forest area loss and a positive trend over time has been shown in some countries and regions such as in Europe where the forest area continued to expand.

Considerable efforts are needed to improve or at least maintain this positive trend. The forest monitoring gives crucial data to decision makers helping them allocate appropriate financial resources for effective forest conservation and management plans. One of these crucial data is Forest Area, which has been selected as one of the 60 indicators for monitoring progress towards the Millennium Development Goals, the 2010 Biodiversity Target and the Global Objectives on Forests. Forest Area and its changes are relatively easy to measure, especially nowadays with the technological progress made in remote sensing fields.

The available high temporal resolution images allow operational and near real-time applications at global, continental and regional scales for forest area mapping and monitoring. The VEGETATION instruments, which were developed with the objective to provide data specifically for vegetation canopy monitoring, are on board two SPOT5 satellites and will be available until 2013. The instruments of this programme (VEGETATION 1 and VEGETATION 2) have monitored and mapped the worldwide vegetation on a daily basis for more than 12 years now. They provide essential information on terrestrial vegetation cover for a large community of users. Therefore, to ensure the continuity of the data after 2012, the SENTINEL mission, which will be launched at the earliest at the end of 2013 (2), will provide data which will fulfil the needs of the current VEGETATION data users. To fill the data gap between VEGETATION-2 and SENTINEL-3, the Belgian Federal Science Policy Office (BELSPO) has decided to build a satellite mission called PROBA-V which is expected to meet all the specifications of the vegetation user community (2). PROBA-V will have an increased spatial resolution in comparison to the VEGETATION instrument, an enhancement expected to provide new opportunities for forest monitoring.

The aim of this study is to evaluate forest cover classification methods on the European Atlantic biogeographical region by applying the Maximum Likelihood (ML), Artificial Neural Network (ANN) and Support Vector Machine (SVM) classification algorithms on simulated PROBA-V, VEGETATION and MODIS data, in order to determine the optimal classification methodology for the region. This work is carried out within the framework of the FM@PROBA-V project[1] towards evaluation and quantification of the improvements of the products for forest monitoring offered by PROBA-V, in relation to its predecessor, VEGETATION. The preliminary results of the study are presented and discussed in this paper.

Study area

The area of interest is the southern part of the Atlantic biogeographical region of Europe (Figure 1) which is located in the north of Portugal. The region is one of the rare mountainous areas in this region, and mountains in the area can reach an altitude of 1500 m. Forest cover is sparse and agriculture is not very intensive in the region (3).

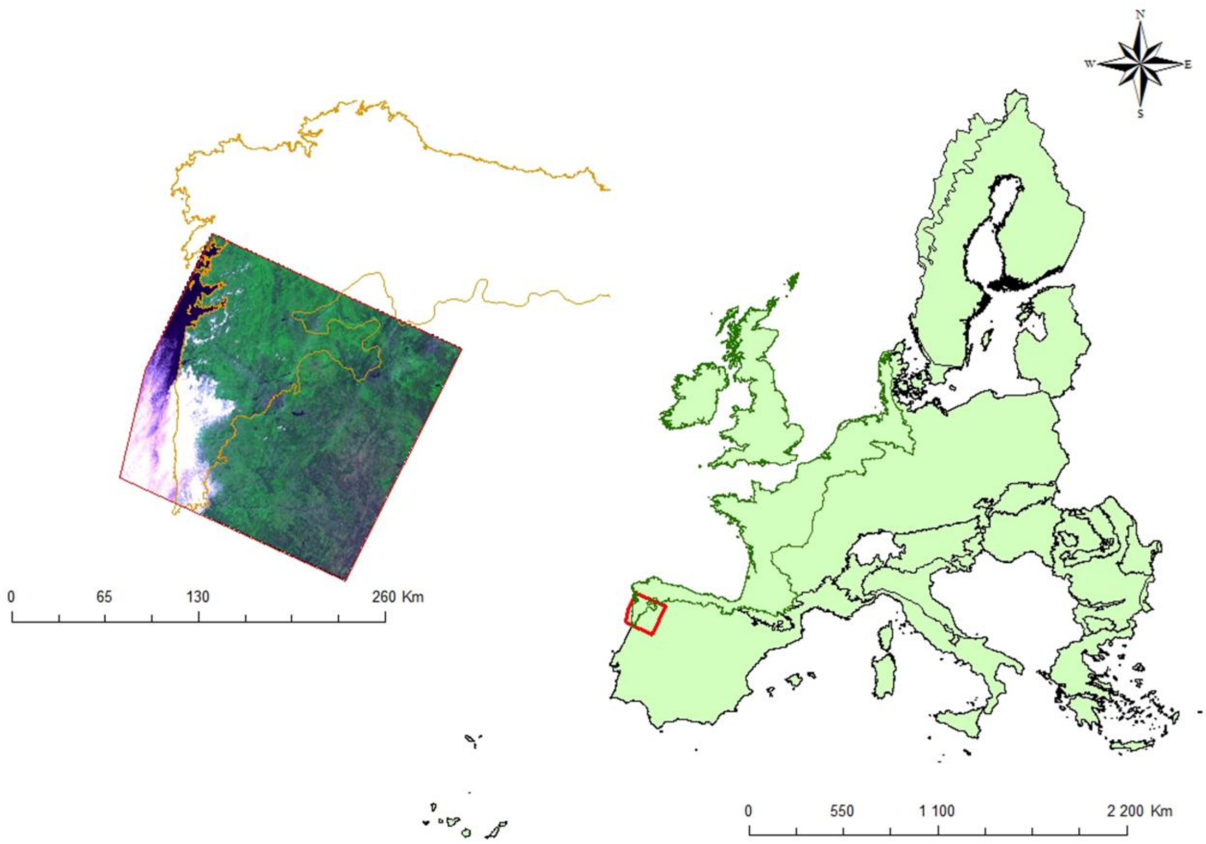

Figure 1: The study area location.

[1] http://fm-probav.maich.gr

METHODS

Dataset and preprocessing

The data used in this study included the images used for the evaluation process of the forest cover mapping in the European Atlantic biogeographical region and the data used to create the reference data for the classifications. The reference data originated from the classification of a Landsat TM image, acquired on 11 July 2009, using training samples and reference points identified from the JRC Forest Cover Map 2006 and Google Earth images. A small subset (20×20 km^2) of that scene was used to simulate a part of a PROBA-V scene, a process undertaken by VITO NV, Belgium. The datasets used to evaluate classification algorithms included the simulated PROBA-V image (Red, NIR, approximated SWIR), a 10-day synthesis VEGETATION (VGT) image at 1000 m spatial resolution with 3 bands (Red, NIR, SWIR), and an 8-day synthesis MODIS image at 250 m spatial resolution (again Red, NIR, resampled/ upgraded SWIR). The simulated PROBA-V subset was relatively small (20×20 km^2) and the results were expected to be of questionable statistical significance. Hence the MODIS image was additionally evaluated, as its similar characteristics with PROBA-V data (Table 1) would allow better evaluation of the potential of the upcoming sensor, and its comparison with VGT data. Since only the red and NIR channels are available at 250 m resolution, the SWIR 500 m channel was resampled to 250 m and two MODIS datasets were evaluated, with and without the SWIR data. All images, including the JRC forest cover map where registered to the Landsat TM image in UTM WGS84, Zone 29N projection, WGS84 datum. The extent of MODIS and VGT images was reduced to a subset matching the extent of Landsat TM, which was topographically normalized with the ASTER GDEM.

Table 1: Image data characteristics

	VEGETATION-2	MODIS	PROBA-V
Blue	0.43 – 0.47	0.45 - 0.47	0.44 – 0.48
Red	0.61 – 0.68	0.62 – 0.67	0.62 – 0.69
NIR	0.78 – 0.89	0.84 - 087	0.79 – 0.90
SWIR	1.58 – 1.75	1.62 –1.65	1.56 – 1.65
Spatial Resolution	1.15 km	250m (R, NIR), 500m (SWIR)	300m (VNIR), 600m (SWIR)

Generation of reference data

High resolution reference data were necessary for this study, in order to train and validate the classification algorithms. However, the only available European-wide data, was the JRC Forest Cover Map, compiled in 2006. Since our datasets were collected in 2009, the Landsat TM data, acquired in 2009, were classified in order to produce a stocked/non-stocked classification map, which would be used to evaluate the classification algorithms in the medium/low resolution datasets.

Classification of Landsat TM

For the classification of Landsat TM in stocked and non-stocked areas, a set of 400 points was generated by stratified random sampling, following the estimation of the extent of the two classes through an unsupervised classification. These points were used as training samples, and were identified as stocked or non-stocked by using the JRC Forest Cover Map and visual interpretation of Google Earth images. (Figure 2). A radius of two pixels was considered around each point in order to identify it as stocked or non-stocked. A second set of 300 points was generated using the same method, to be used for the validation of the classification.

A Maximum Likelihood supervised classification was performed on three different band combinations: 6 bands (excluding the thermal band), 4 bands (Blue, Red, NIR and SWIR), and 4 bands + NDVI generated from the Landsat TM data, in order to find the best result. The same training samples were used for the three band combinations. The four classes defined in these classifications were "stocked areas", "non-stocked areas", “water” and "No data". The confusion matrices, for the three classification’s results, were performed using the same 300 validation points.

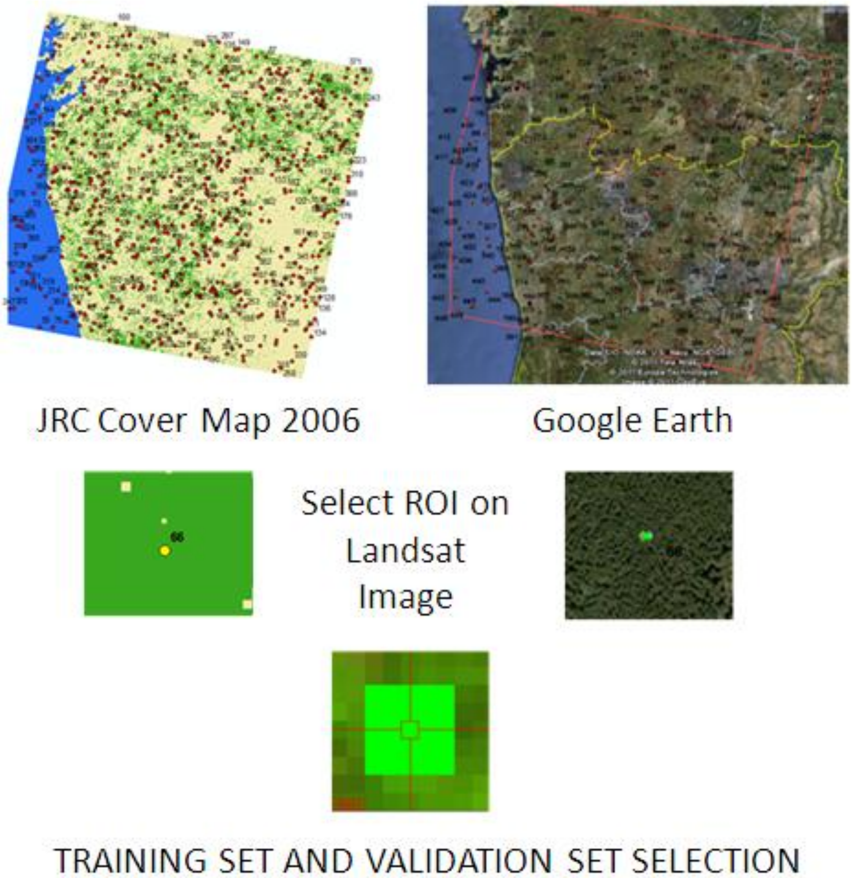

Figure 2: Training samples and validations point selection (ROI: Region of Interest).

Forest cover reference maps creation

The classification map with the higher accuracy was resampled to a 10 m resolution, using nearest neighbour, and the “no data” and “water” classes were masked out, leaving the two main classes, “stocked” and “non-stocked” with values of 1 and 0, respectively.

In order to create reference maps with the same resolution as the medium/low resolution images, the 10 m resolution map was aggregated to 250, 300 and 1000 m resolution. The new value of each aggregated pixel was the percentage of the original “stocked” pixels that were aggregated to the lower resolution pixel. According to the pixel values obtained after aggregation, each map was reclassified to five classes according to the percentage of forest cover (0 – 10%, 11% - 30%, 31% - 50%, 51% - 75%, 76% - 100%), see Figure 3. Those classes were chosen in order to ensure that the entire range of forest cover fractions would be represented within the broad “stocked” and “non-stocked” classes.

The Maximum Likelihood supervised classification on the three different combinations of Landsat TM bands produced the highest overall accuracy (93.12%) and Kappa Coefficient (0.904) when all six bands (excluding the thermal) were used, as shown in Table 2. The resulting stocked/non-stocked map was used as a reference map for the study.

Table 2: Results of the accuracy assessment of Landsat TM classification

Image	TM 6 bands	TM 4 bands	TM4 + NDVI
Overall accuracy	93.1214%	92.9878%	92.4135%
Kappa coefficient	0.9045	0.9027	0.8947

Evaluation of classification algorithms

The produced reference data were used to generate training samples for the classifications and the validation points for the confusion matrices. Training samples were generated for the five classes using stratified random sampling method and subsequently the regions of interests were merged in order to create the two main classes, Non-stocked (0% - 10% and 11% - 30%) and Stocked (31% - 50%, 51% - 75%, and 76% - 100%).These training samples were used to perform the Maximum Likelihood (ML), Support Vector Machine (SVM) and Artificial Neural Network (ANN) supervised classifications on the two- and three-band MODIS, VGT and PROBA-V datasets. Confusion matrices were generated for all classification results using the LandsatTM-derived 250m, 300m and 1000m reference maps , respectively for MODIS, PROBA-V and VGT classifications.

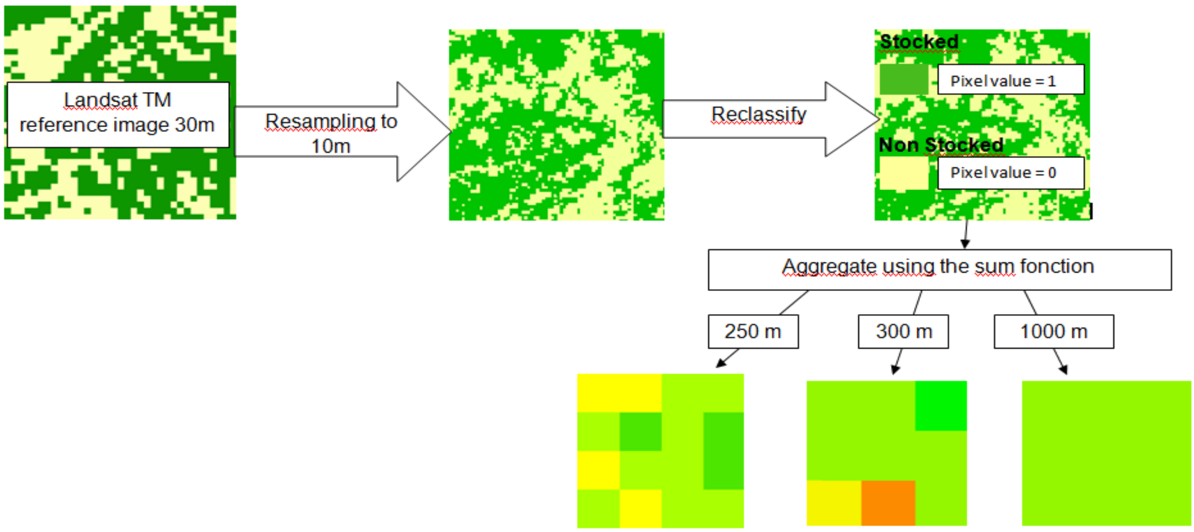

Figure 3: Flowchart of reference maps creation.

RESULTS

The results of the confusion matrices are presented in Figure 4. It appeared that the two-band 250 m MODIS dataset produced more accurate classifications with all classification methods, in comparison with the three-band dataset, which employed the SWIR data resampled from 500 to 250 metres.

Comparison between the classification methods for all datasets shows that the ANN classification gives a slightly higher overall accuracy than the other methods for the two- and three-band MODIS data, as well as the simulated PROBA-V, while SVM gives the highest overall accuracy for VGT. However the differences between the different classifiers for each dataset are very small and could be considered to be of low significance.

According to these preliminary results, the best accuracy and kappa coefficients were achieved by the ANN classification on MODIS image with two bands at 250 m spatial resolution and the lowest overall accuracy was performed by ML classification on simulated PROBA-V image (Figure 4).

In addition to the classification accuracy the percentage of stocked area distribution in each map was also calculated (Figures 5 and 6). For MODIS and VGT data, the estimated forest cover was overestimated significantly, from 27% which was the estimated forest cover using the Landsat TM data, to 41 to 48%. On the contrary forest cover estimation with PROBA-V, using the small subset of the scene, for which the simulated PROBA-V data were available, showed significant differences between the classifiers. The Landsat TM data produced an approximate 70% forest cover, while Maximum Likelihood underestimated the forest cover (57.55%) and MODIS and VGT overestimated the forest cover (80.72 and 79.39% respectively).

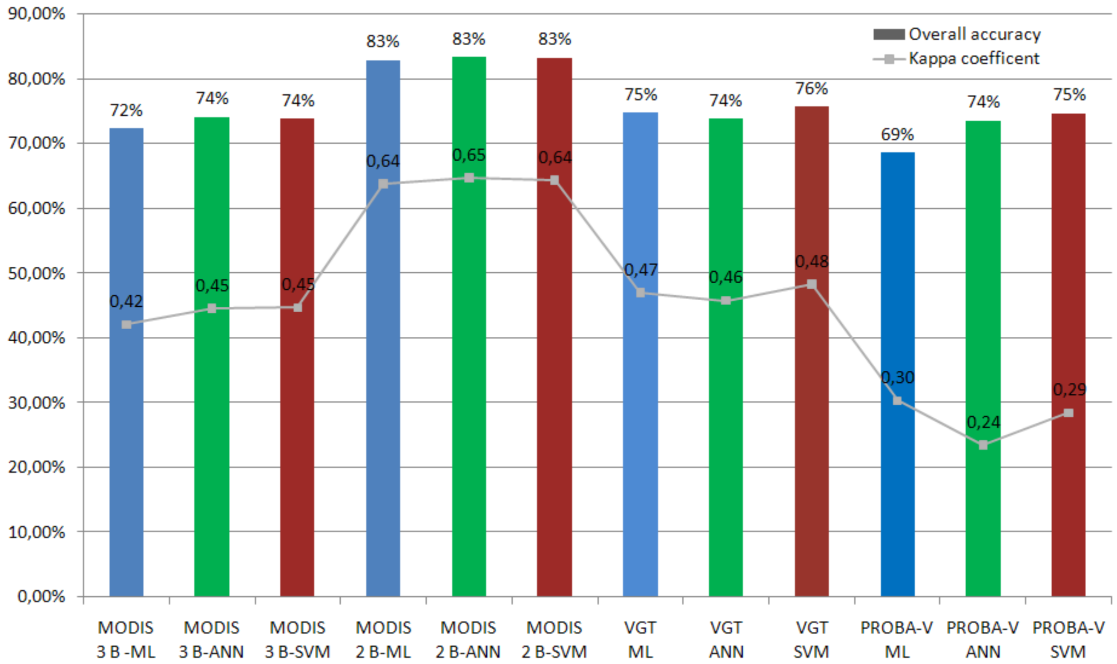

Figure 4: Overall accuracy and Kappa Coefficients or ML, ANN and SVM classification result on MODIS, VGT and PROBA-V (B: bands).

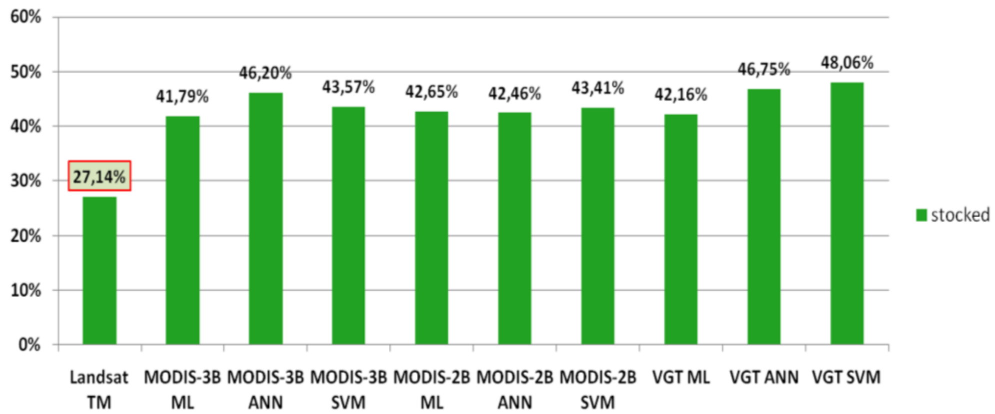

Figure 5: Class distribution of the different classification results comparing to Landsat TM reference (B: bands).

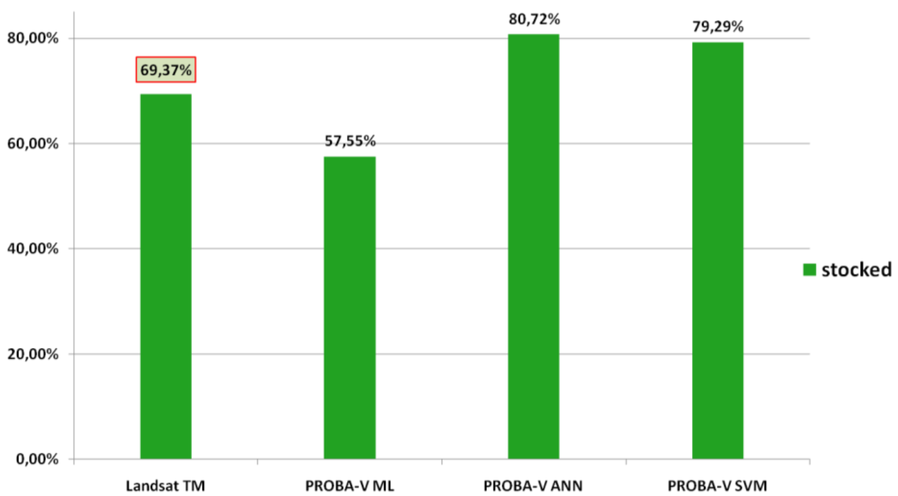

Figure 6: Classes distribution of simulated PROBA-V classification compared to the subset of the Landsat reference map.

Discussion and Conclusions

The error matrix is the most common method to evaluate classification accuracy (4), and it is the starting point for many analysis techniques (4,5). In addition to the overall accuracy, the entire confusion matrix was considered through the use of the kappa statistic. These analyses were used to allow a comparison between classification results.

Forest cover in the study area showed significant fragmentation, which tends to pose problems when classifying satellite images (6). This fact proved to be detrimental to the performance of certain datasets. The main aim of this study was to compare the potential performance of PROBA-V in estimating forest cover, with that of VEGETATION (VGT). Because of the small extent of the simulated PROBA-V scene, 250 m MODIS data were also evaluated as an alternative sensor with comparable spatial resolution, which could deliver more statistically significant results. The lack of 250 m SWIR data (the native 250 m product contains data only in the Red and NIR channels) forced the resampling of the 500 m SWIR data to 250 m. Comparison between the classification accuracies between the three-band (with the resampled SWIR data) and the two-band (without the SWIR data) MODIS data showed that the latter achieves more accurate classifications. Spectral variation of the SWIR signal within the 500 m pixel could not be retrieved with the resampling process and, as a result, the combination of accurate 250 m Red and NIR data with "false" SWIR data led to considerable misclassifications between stocked and non-stocked areas On the contrary, the use of just the Red and NIR data produced more accurate classifications.

Comparison between classifications of the two-band MODIS data and the VGT data, revealed that the former again produced more accurate classifications. The availability of SWIR data in the VGT dataset could not assist the classification sufficiently in order to make up for the increased uncertainty brought about by the 1 km spatial resolution of the data. As expected, extraction of forest cover information in a fragmented landscape using low resolution data, proves to be quite problematic at a local to regional scale (4).

The simulated PROBA-V data covered only a 20×20 km^2 area and consisted of approximately 3600 pixels. From a statistical point of view it is very difficult to extract statistically significant results from such a small sample size, which is the reason MODIS data were also used in this study. Nevertheless, the classification accuracy of PROBA-V data was very similar to the one produced by VEGETATION. On the other hand, the low kappa values of the classifications are caused by the amplification of the errors in the confusion matrix, brought about the small sample size.

The accuracies of the classifications produced by the different classification algorithms on the same dataset showed that there is effectively no difference in their performance on all three datasets. The only exception was Maximum Likelihood classifier on the simulated PROBA-V data, which was by 5% less accurate in comparison with ANN and SVM, a fact probably attributed to the quality of the simulated PROBA-V data The Artificial Neural Network classification provides the most accurate classifications and appears to deal best with the mixture of different cover types within each coarse resolution pixel (7) and the fragmentation of the landscape (8).

Non-thematic errors may be one major problem in the use of confusion matrix and associated accuracy analysis (9). This is particularly true for errors of image misregistration, since the images were registered to an image of a higher spatial resolution. The methodology used of aggregation and forest coverage percentage class was used to minimize these errors.

In conclusion, comparison between classification algorithms on MODIS and VGT data, showed no significant differences between them, in terms of classification accuracy. The higher accuracy achieved by those classifiers when using the MODIS data, which have comparable spatial resolution with PROBA-V, compared to that achieved with the VGT data, suggest that the increased spatial resolution will provide more accurate forest cover mapping, particularly in areas with fragmented forest coverage. The usefulness of SWIR data could not be evaluated, but it could prove useful in discriminating between forest and agricultural land, and assist even further the accurate mapping of forest cover. More research on this topic is recommended in order to quantify the usefulness of SWIR data at a 300 m resolution.

ACKNOWLEDGEMENTS

This study was funded by the Belgian Federal Science Policy Office (BELSPO) within the framework of the Preparatory Programme of the PRoject for On-Board Autonomy – VEGETATION (PROBA-V), and carried out as part of the research of the *Forest Cover Change Monitoring with PROBA-V: Potentials and Limitations on European Terrain (*FM@PROBA-V) project.

References

1 Food and Agriculture Organization of the United Nations (FAO), 2010. Global Forest Resources Assessment 2010. Main Report, 378 pp. Executive Summary.

2 Schyns J C, 2010. The PROBA-V Mission: Data continuity for daily global vegetation monitoring. 30th EARSeL Symposium Remote Sensing for Science, Education and Culture, p. 635-640

3 Condé S, D Richard, N Liamine, A S Leclère, D Moss & H O Hansen, 2002. The Atlantic Region. 60pp. In: Europe's Biodiversity - Biogeographical Regions in Europe (European Environment Agency) EEA Report 1/2002

4 Lunetta R S, R G Congalton, L K Fenstermaker, J R Jensen, K C McGwire & L R Tinney, 1991. Remote Sensing and Geographic Information system data integration: Error sources and research issues. Photogrammetric Engineering and Remote Sensing, 57(6): 677-687

5 Congalton G R & K Green, 2009. Assessing the accuracy of remotely sensed data: Principal and practice. 2nd Edition, p. 105-106 (CRC Press, Taylor & Francis) 193 pp.

6 Lin H, J Wang, Y Bo & J Bian, 2008. The impact of landscape patterns on the accuracy of remotely sensed data classification. In: Proceedings of 8th International Symposium on Special Accuracy Assessment in Natural Resources and Environmental Sciences (Shanghai, China, June 25-27, 2008) pp. 219-225 (last date accessed: 21 Nov 2011)

7 Moody A, S Gopal & A H Strahler, 1996. Artificial neural network response to mixed pixels in coarse-resolution satellite data. Remote Sensing of Environment, 58: 329-343

8 Moreira F, F C Rego & P G Ferreira, 2001. Temporal (1958–1995) pattern of change in a cultural landscape of northwestern Portugal: implications for fire occurrence. Landscape Ecology, 16: 557-567

9 Strahler A H, L Boschetti, G M Foody, M A Friedl, M C Hansen, M Herold, P Mayaux, J T Morisette, S V Stehman & C E Woodcock, 2006. Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps, 10 (Office for Official Publication of the European Communities) 43 pp.